~~200~~ Repl-
UL

Psychological Monographs

General and Applied

492
60
Fairweather, Simon, Gebhard, Weingarten,
74
5

Relative Effectiveness of Psychotherapeutic Programs: A Multicriteria Comparison of Four Programs for Three Different Patient Groups

By

G. W. Fairweather, R. Simon, M. E. Gebhard, E. Weingarten
J. L. Holland, R. Sanders, G. B. Stone, and J. E. Reahl
Veterans Administration Hospital, Perry Point, Maryland

Price $1.00

Edited by Norman L. Munn
Published by the American Psychological Association, Inc.

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

NORMAN L. MUNN, Editor

Department of Psychology, Bowdoin College
Brunswick, Maine

Consulting Editors

ANNE ANASTASI
FRANK A. BEACH
ARNOLD M. BINDER
W. J. BROGDEN
ROBERT R. BUSH
JOHN F. DASHIELL
JAMES J. GIBSON
D. O. HEBB
EDNA HEIDBREDER
FRANCIS W. IRWIN
JAMES J. JENKINS
HAROLD E. JONES
DANIEL KATZ
BOYD MCCANDLESS
DONALD W. MACKINNON
QUINN MCNEMAR
LORRIN A. RIGGS
CARL R. ROGERS
RICHARD L. SOLOMON
ROSS STAGNER

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual*. Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual*. Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

ARTHUR C. HOFFMAN, Managing Ed.; HELEN ORR, Promotion Mgr.; JANET LEBAN, Editorial Asst.

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

COPYRIGHT, 1960, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Vol. 74, No. 5 Whole No. 492, 1960

Psychological Monographs: General and Applied

RELATIVE EFFECTIVENESS OF PSYCHOTHERAPEUTIC PROGRAMS:

A MULTICRITERIA COMPARISON OF FOUR PROGRAMS FOR THREE DIFFERENT PATIENT GROUPS[1]

G. W. FAIRWEATHER,[2] R. SIMON,[3] M. E. GEBHARD, E. WEINGARTEN, J. L. HOLLAND,[4] R. SANDERS,[5] G. B. STONE, and J. E. REAHL

Veterans Administration Hospital, Perry Point, Maryland

PSYCHOTHERAPISTS in neuropsychiatric hospitals are continually faced with the problem of formulating feasible and economical psychotherapeutic programs for a heterogeneous population of patients. Historically, such programs have usually been established upon the theoretical propositions accepted by those in positions of authority and, frequently, little attempt has been made to evaluate their effectiveness. In addition, when such evaluations have been carried out, a single criterion is selected and its relationship to other critieria has been ignored. The present study attempts to compare the relative efficiency of four common psychotherapeutic programs for three patient groups using a number of diverse inventories, projective devices, and rating scales.

Probably the most common treatment program in neuropsychiatric hospitals today provides each patient with an individual work assignment and planning for post-hospital adjustment. This program is based upon the assumption that the hospital work situation can be therapeutic. For that reason, the hospital personnel attempt to place the patient in a work situation suited to his unique skills and needs. This opportunity for maximal use of his abilities in a job setting may enhance the patient's self-esteem and result in more adaptive behavior in the extramural setting.

A second common treatment program employs individual work assignments, individual psychotherapy, and departure planning. This assumes that underlying conflicts must be resolved before any effective personality change can occur and, further, that personality change is essential for extramural adjustment. It assumes too that the relationship between the psychotherapist and patient is the most effective instrument for resolving unconscious conflicts. A host of publications and documents signify that individual psychotherapy has, indeed, become one of the most common psychotherapeutic approaches of our time.

A third treatment program uses individual work assignments, group psychotherapy, and departure planning. It differs from the two preceding programs only with respect

[1] The authors wish to express their appreciation to the entire staff of VA hospital, Perry Point, Maryland, for their assistance in making this study possible. We are particularly indebted to Sarah Shifflet, Anna Wirsing, Grace Pruitt, and James Sprouse, nursing assistants, and Mary McMenamin, nurse, who served on the experimental ward. For clerical and statistical help we are grateful to Ralph Forsyth, Barbara Creswell, Kathryn Boyd, Itsuye Sakai, L. Arlene Zealor, Marian Barry, and Lucy Yoshizawa. To Helen Jefferys, Marge Edelson, and John Webster we express gratitude for their social work services. We are especially indebted to Earl Brannon, Manager, and Willy Oppler, Chief of Professional Services, for their administrative assistance in carrying out this project. We are grateful to Roy E. Buehler for his contributions to the theoretical aspect and Laura Kent for her aid in preparation of the manuscript.

[2] Now with VA Hospital, Palo Alto, California.

[3] Now with National Institutes of Health, Bethesda, Maryland.

[4] Now with National Merit Scholarship Corporation, Evanston, Illinois.

[5] Now with Philadelphia State Hospital, Philadelphia, Pennsylvania.

to group psychotherapy. Proponents of group psychotherapy, as the medium for personality change, maintain that, since man has learned his maladaptive responses in a social milieu, the most effective vehicle for changing such responses is group psychotherapy. In this program the situation for personality change is the pattern of interpersonal relationships occurring in group therapy.

The fourth program, albeit used less frequently than the preceding three, is a social interaction, or "group living therapy." This rests upon the proposition that the total life situation of the patient can be used as an instrumentality for personality change. Therefore, the patients live and work together, and, as an integral part of this program, participate in group psychotherapy. This program differs from the other three in that all its functions are group oriented. The rationale is that changes in maladaptive perception and subsequent behaviors are best effected by planned social situations which provide an opportunity for correcting inappropriate interpersonal responses. Each situation provides an atmosphere for group decision and responsibility conducive to group cohesiveness. It is the cohesiveness of the group which mirrors the accepted norms of the society and provides the background for corrective emotional experiences. In this atmosphere, all group activities are believed to be potentially therapeutic.

In general, psychotherapeutic programs such as those described have been more successful with neurotic and acute psychotic patients than with chronic psychotic patients.

From the aforementioned experience and rationale two general hypotheses evolve. They are:

1. Treatment programs involving psychotherapy effect more change in the direction of adaptability than work programs alone.

2. Nonpsychotics and short-term psychotics change more in the direction of greater adaptability with any treatment program than long-term psychotics.

PROCEDURE

The experiment took place on a ward in a large Veterans Administration neuropsychiatric hospital. The number of patients on the ward at a time varied from 40 to 62. The ward staff included a psychiatrist, two psychologists, a social worker, a psychological intern, a nurse, and four nursing assistants. As for its physical features, there were three large dormitories, a recreation room, a reading and visitors room, and a television room. All patients in the study remained on this ward for the duration of their treatment.

The general ward program for all patients in the experiment included responsibility for ward housekeeping and, in addition, two recreational evenings per week. The first of these took place on Sunday or Monday and included a structured recreational situation. During these periods a member of the special services division of the hospital, along with a group of women volunteers from the community, planned recreational situations in advance, i.e., dances, table games, etc. The second recreational situation, on Thursday evenings, was an unstructured situation; the women volunteers arrived on the ward and, together with the patients, planned recreational activities for the evening. Usually informal groups were formed spontaneously. They played cards, danced, and engaged in similar activities. In addition, the ward recreation rooms were open from 11 to 12 in the morning and from 3 to 9 o'clock in the evening. When the patients were not participating in their specific treatment programs, they could watch TV, play pool or cards, read, and engage in other activities. During this time a nursing assistant was on duty.

The Four Experimental Conditions

In the first treatment program, a rehabilitation team[6] provided each patient with an individual work assignment and planning for posthospital

[6] Consisting of a psychologist, psychiatrist, counseling psychologist, social worker, and psychological intern. Any further use of the term "rehabilitation team" refers to these individuals.

living. This group constitutes the control group for the study and will be referred to as Group C. For the second group (Group I), the treatment program consisted of an individual work assignment, individual psychotherapy, and departure planning outlined by the rehabilitation team. The treatment program for the third group (Group G) included an individual work assignment, group psychotherapy, and departure planning by the rehabilitation team. For the fourth group (Group GG), the treatment consisted of a group work situation, group living situation, and group psychotherapy.

Upon their arrival on the ward, members of the C, G, and I groups were interviewed by a counseling psychologist and given individual work assignments. In addition, the G and I groups were given therapy appointments. Groups C, G, and I were responsible for ward housekeeping in all dormitories except the one occupied by the GG group. For these three groups, each patient was assigned to ward housekeeping for a period of two weeks. This was done on a rotating basis.

The GG group lived together in a single ward dormitory. Their work assignment was to produce a hospital newspaper every two weeks. Each afternoon at 1 o'clock they reported to a room which had been set aside as the newspaper workroom. Here, under the supervision of a staff member, they had the total responsibility for running the newspaper: they made decisions about the content of the paper, the reporting of the articles, its reproduction, and all other matters relating to the administration and design of the paper. The first hour of each afternoon was devoted to a discussion of their work. In addition, they received group therapy and were responsible as a group for the housekeeping in their dormitory.

Therapists and Therapy

Two staff psychologists served as therapists for the patients in the G and GG groups. One of these had four years of postdoctoral psychotherapeutic experience and the other had two years. The 48 patients in these two groups were divided evenly between the therapists.

The individual therapy patients (Group I) were carried by 15 psychotherapists with a wide range of experience. Five clinical psychology staff members, whose range of experience varied from two to five postdoctoral years, provided individual psychotherapy. One staff psychiatrist, with six years of psychotherapeutic experience, carried an individual therapy patient. In addition, four clinical psychology trainees and four psychiatric residents provided individual psychotherapy. All were supervised by consultants.

Both the group psychotherapy and individual psychotherapy sessions were psychoanalytically oriented. Group sessions were held for 1—1/2 hours two times weekly and individual therapy sessions varied from two to four per week depending upon the judgment of the psychotherapist. Groups averaged six patients and ranged from four to eight. The groups were open groups and usually consisted of neurotics, acute psychotics, and chronic psychotics.

Criterion for Discharge or Failure

At the weekly meeting of the rehabilitation team, each patient's current hospital adjustment was discussed and his readiness for discharge was evaluated. No patient left the treatment program unless it had been unanimously agreed that he was ready for discharge. In addition to the 85 patients discharged, there were 11 who were considered failures. The criterion for failure was the team's decision that the patient had shown no progress, or had shown a decrement in his general adjustment after six months of treatment. Such patients were considered failures and were sent from the experimental ward to other hospital wards where a more appropriate treatment program was possible.

Sampling Procedures

When the program started, each patient coming to the ward was randomly assigned to one of the four treatment groups. After a sufficient sample in each group had been accumulated by random assignment, new patients were placed by matching across treatments on age, diagnosis, and length of hospitalization. When new patients who did not meet the matching criteria arrived on the ward, they were randomly assigned to one of the four treatment groups. Thus, over a period of time a treatments by diagnoses factorial analysis of variance design was attained with 24 patients in each of the four treatment groups and 32 patients in each of the three diagnostic groups.

The four treatments have been discussed. The three diagnostic groups were: (*a*) nonpsychotic, (*b*) short-term psychotic, and (*c*) long-term

psychotic.[7] Each patient in the nonpsychotic group had been diagnosed a nonpsychotic. Each patient in the short-term psychotic group had less than one year of previous hospitalization in addition to his psychotic diagnosis. Each patient in the long-term psychotic group had a psychotic diagnosis and previous hospitalization exceeding one year.

A descriptive comparison of the treatment and diagnostic groups appears in Table 1 and 2, respectively.

Table 1 indicates that the treatment groups are quite homogeneous with respect to diagnosis, amount of drugs received, age, education, race, length of previous hospitalization, pensions for NP illness, age of first hospitalization, employment history, occupational status, and marital status.

On the other hand, Table 2 reveals significant difference among the three diagnostic groups. The stratifying variable of previous weeks of NP hospitalization significantly differentiates the three diagnostic groups as do the pension, employment, and marital status variables. No significant differences are found for age, education, race, age of first NP hospitalization, weekly salary, occupational status, or percentage of patients on tranquilizing medications. Viewed collectively, the differences suggest that the long-term psychotic is more maladaptive than either the nonpsychotics or short-term psychotics.

[7] Diagnostic groups are listed in the tables as follows: nonpsychotics (N), short-term psychotics (Ps), and long-term psychotics (Pl).

Experimental Instruments and Testing Situation

In selecting criteria for evaluating change, the experimenters attempted to include measures that were frequently used for this purpose and encompassed a broad range of behavior. The areas selected for assessment included individual behavior, self-evaluation, fantasy, and follow-up. Specifically the instruments were: Ward Behavior, MMPI, Q sort, Holland Vocational Preference Inventory (HVPI), TAT and a Follow-up Rating

TABLE 1

COMPARISON OF THE FOUR TREATMENT GROUPS ON DEMOGRAPHIC CHARACTERISTICS

Variable	GG	G	I	C	Test of Significance
Matching Variables					
Diagnoses					
% character disorders	12.5	12.5	12.5	12.5	
% neurotics	20.8	20.8	20.8	20.8	$\chi^2 = 0$
% psychotics	66.7	66.7	66.7	66.7	
Mean age	36.25	33.17	34.21	34.92	$F = \ldots$ [a]
Mean weeks of prior NP hospitalization	51.6	49.2	53.8	54.8	$F = \ldots$
Concurrent Variable					
% receiving tranquilizing drugs	54	58	58	42	$\chi^2 = 1.58$
Demographic Variables					
Education					
Mean grade completed	10.5	10.6	11.0	11.2	$F = \ldots$
Race					
% white	79	75	75	75	
% Negro	21	25	25	25	$\chi^2 = .17$
% pensioned for NP illness	67	50	67	58	$\chi^2 = 1.52$
Mean age of first NP hospitalization	28.5	27	29	31	$F = \ldots$
Employment					
Mean % time employed in past 10 years	71.5	75	68	71	$F = \ldots$
Mean weekly salary of last job	74.3	63.2	68.5	62.7	$F = \ldots$
Occupational status					
% professional-managerial	12.5	20.8	25	16.7	
% skilled trades, sales, clerical	33.3	20.8	20.8	20.8	$\chi^2 = 2.48$
% unskilled workers	54.2	58.4	54.2	62.5	
Marital Status					
% currently married	25	33	25	25	
% single or divorced	75	67	75	75	$\chi^2 = .53$

[a] The F ratio value . . . indicates that the within-group variance was greater than the between-group variance.

TABLE 2

COMPARISON OF THE THREE DIAGNOSTIC GROUPS ON DEMOGRAPHIC CHARACTERISTICS

Variable	N	PS	PL	Test of Significance
Matching Variables				
Mean age	36.09	34.41	33.31	F = . . .[a]
Mean weeks of prior NP hospitalization	25.34	20.47	111.22	F = 45.91**
Concurrent Variables				
% receiving tranquilizing drugs	37.5	56.3	65.6	χ^2 = 5.27
Demographic Variables				
Education				
Mean grade completed	11.03	10.34	11.09	F = . . .
Race				
% white	81.3	65.6	81.3	χ^2 = 3.94
% Negro	18.7	34.4	18.7	
% pensioned for NP illness	37.5	59.4	90.6	χ^2 = 19.46**
Mean age of first NP hospitalization	31.19	29.75	25.88	F = 3.06
Employment				
Mean % time employed in past 10 years	79.22	83.50	51.09	F = 17.02**
Mean weekly salary of last job	68.06	73.72	56.13	F = 2.37
Occupational status				
% professional-managerial	12.5	21.9	21.9	
% skilled trades, sales, clerical	34.4	18.8	18.8	χ^2 = 3.32
% unskilled workers	53.1	59.3	59.3	
Marital status				
% currently married	28.1	46.9	6.2	χ^2 = 13.39*
% single or divorced	71.9	53.1	93.8	

[a] The F ratio value . . . indicates that the within-group variance was greater than the between-group variance.
*.01 level of significance.
**.001 level of significance.

Scale.[8] Their construction and scoring will be discussed in subsequent sections devoted to the findings obtained with each assessment device.

Every patient in the study was tested twice with the instruments listed above. The first assessment was made at the time the patient transferred to the experimental ward and the second shortly before he left the ward. The tests were administered during scheduled sessions in an office building which was located some distance from the ward. The same examiner administered both initial and terminal tests. The initial testing occurred, with a few exceptions, on the first Monday afternoon and Tuesday morning following a patient's transfer to the experimental ward. At that time the patient was instructed by the ward nurse to report to the office building "for testing." In a few cases, due to uncontrollable circumstances such as illness, the time interval between transfer to the ward and testing was somewhat longer, though it never exceeded two weeks. The second administration usually was scheduled for the Monday afternoon and Tuesday morning of the week during which the patient was to leave the experimental program. Here again there were occasional exceptions.

The tests were administered to groups which averaged three or four patients and ranged from two to nine. Groups undergoing initial and terminal testing worked in different rooms.

The following instructions were given for the initial testing:

> All of you have recently been transferred to Ward 19A from other parts of the hospital. Part of the routine procedure connected with transfer to 19A is some psychological testing and this is why you were asked to come over here today. The purpose of these tests is to provide information about you which will be of help to your doctors on the ward in helping you. It is therefore in your interest to cooperate to the best of your ability in doing these things which you will be asked to do, and I hope you

[8] The Ward Behavior Scale, Q sort, and Follow-up Scale have been deposited with the American Documentation Institute. Order Document No. 6239 from the A.D.I. Auxiliary Publication Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $2.50 for photoprints or $1.75 for 35 mm. microfilm. Make checks payable to Chief, Photoduplication Service, Library of Congress.

will do so. Now, some of you may have done some of these things before, either when you first came to Perry Point, or at some other time. But, even so, even if you have done them only a few days ago we would like you to do them again, since it is important that we get the most accurate picture that we can of you as you are now, as you are entering 19A. There are several things which you will be asked to do and these will take, probably, all of the afternoon and part of tomorrow morning. Unless there are any questions, we will begin.

The examiner then distributed the Q sort cards and place cards, drawing the most-like-6 and least-like-6 cards out of the stack and placing them separately in front of the patient. He then proceeded with the instructions as follows:

I have just placed before each of you a stack of cards, each of which has a typewritten statement on it. What I would like you to do is to sort these cards according to the way the statements apply or don't apply to you. The way I want you to start is to go through the whole stack, reading all the statements. Out of these cards you are to select the six which are most like you the way you are, which are most true of you, and to put them on the card which says most-like-6. Also select the six statements which are least like you the way you are, which are least true of you, and put them on the card marked least-like-6. Now remember, in doing this you are supposed to go through the whole stack of cards.

After a patient had completed the first sort, the examiner counted the cards he had placed in each pile, attached them and placed them in a separate envelope for each patient. Then he drew the most-like-9 and least-like-9 cards from the stack and said:

Now go through the remainder of the cards and, out of all the cards you have left, select the nine statements which are most like you and the nine which are least like you and put them on these cards. Again, go through all the cards you have left.

This sorting was repeated twice more with 15 and 20 cards each.

After a patient had completed the Q sort and the examiner had checked the numbers in each pile, an HVPI booklet and answer sheet were placed before him and the examiner said:

The next thing you will do is to go through this list of jobs or occupations and indicate on the answer sheet your first impression as to whether a particular job or occupation appeals to you. If it does appeal to you, black in between the dotted lines under the column with the heading A. A stands for appeal. If it does not, black it in under heading D. Just go on the basis of your first impression, without spending much time thinking about any one item. Put your name in this space before you begin.

After a patient had completed this test the examiner put before him an MMPI booklet and answer sheet saying:

This is the last thing which we will do this afternoon. This booklet contains a number of statements and I want you to indicate whether you believe them to be true or false of you by making a mark under the column headed T if true, F if false.

At the end of the afternoon session, the patients were requested to return the following morning. At that time the examiner again distributed the Q sort material and said:

You remember that the first thing we did yesterday afternoon was to sort these cards into different piles according to whether the statements on them applied to you or didn't. This morning I would like you to sort them again but on a different principle. Instead of sorting them according to the way you are, you should sort them according to the way you would like to be. This means that, in the first step, you should pick out the six statements that you would most like to be like, that you would most like to be true of you, regardless of whether they are actually true of you or not, and the six that you would least like to be true of you regardless of whether they are true of you or not. For example, let's assume that one of the statements is, "I have a million dollars." Yesterday you would have put that in one of the least-like piles since that isn't true of you. Today you might possibly put it in one of the most-like piles since, although it isn't true of you, you might wish that it were true of you. Now, are you all clear on this difference between today's and yesterday's sort? Again, you are to go through the whole stack each time you make a sort.

After completion of the initial step the patients were asked to sort 9, 15, and 20 statements each into most-like and least-like piles. The procedure was the same as in the initial sort.

The examiner then distributed the TAT material and said:

This is the last thing you will be asked to do. I have given you each a series of pictures which show one or more persons in some kind of a setting. What you are to do in this is to write a story about each picture, a story which has in it the answers to these questions: What might be happening? What might be going on? What might the people be thinking and feeling? What might have led up to this? What might the outcome be? You should not spend more than five minutes on any one story after you have thought of something. Please put your name on each sheet of paper you use.

After a patient had completed the TAT he was dismissed.

The instructions for the second administration varied, depending upon whether the group to be examined consisted entirely of prospective dis-

charges or whether it included "failures" who were about to be transferred to other wards. In the former case the testing session was introduced as follows:

> You may remember that when you first came to 19A you came here and were given several tests. These were part of the routine procedure connected with transfer to that ward. Now you are about to leave the hospital and you will be asked to take these tests again. They are part of the routine procedure connected with the discharge process. Again I would like you to cooperate to the best of your ability.

If the group included "failures" the instructions were:

> When you first came to 19A you came here and were given several tests. These were part of the routine procedure connected with transfer to that ward. It is part of our procedure also, that these tests be taken twice, once upon entering the ward and, the second time, after you have been on the ward a certain length of time. This will be the last time you will be asked to take them. Again I would like you to cooperate to the best of your ability.

In other respects the procedure was the same as that in the initial administrations.

All initial and terminal test data were filed following their completion. The experimenters did not see the data until after the last subject had completed the program.

Attrition Group

During the 24-month period covering the course of the experiment, 25 patients were dropped from the experimental program. These failed to complete the program for one of the following reasons: (*a*) required major surgery; (*b*) discharged with less than 21 days in the experimental program; (*c*) left the hospital without permission; (*d*) developed more acute psychotic symptoms requiring transfer to a closed ward; (*e*) refused to cooperate in posttesting. Table 3 summarizes the reasons for attrition by experimental group.

TABLE 3

REASONS FOR ATTRITION BY EXPERIMENTAL GROUP

	GG	G	I	C
Required major surgery		1		2
Discharged in less than 21 days	4		1	2
Developed more acute psychotic symptoms	1		2	3
Left without permission	2	2	1	1
Refused posttesting		1		2
Total	7	4	4	10

TABLE 4

MEAN DAYS IN TREATMENT

	GG	G	I	C	Mean
N	119.5	65.4	117.8	86.0	97.18
Ps	78.6	61.4	129.8	89.5	89.83
Pl	172.3	163.9	191.1	92.5	154.95
Mean	123.47	96.90	146.23	89.33	

Time in Treatment

Since the purpose of this investigation is to determine the efficacy and economy of the various treatment programs for different diagnostic groups, the amount of time spent in treatment, an important consideration in evaluating the efficacy and economy of treatment, was allowed to vary. Table 4 presents the mean time, in days, for the different groups and Table 5 shows the analysis of variance for the same data. Table 5 indicates that significant differences obtain for treatments and diagnoses but not for their interaction.

Table 4 reveals that patients who received individual psychotherapy remained in the hospital longer than the other three treatment groups. A further glance at Table 4 shows that the highly significant difference

TABLE 5

ANALYSIS OF VARIANCE OF TIME IN TREATMENT

Source of Variation	*df*	Mean Square	*F*
Treatments	3	16,228.37	3.98*
Diagnoses	2	40,714.00	9.99**
Interaction: Treatment × Diagnoses	6	5,451.67	1.34
Within groups	84	4,073.69	
Total	95		

* Significant at .05 level.
** Significant at .001 level.

between diagnostic groups results from the much longer treatment period of the long-term psychotic group when compared with the other two diagnostic groups. Table 4 also reveals that the long-term psychotic control group remained in treatment a shorter period of time than long-term psychotics in psychotherapy.

RESULTS

The results for each experimental instrument are presented in the following sections.[9] Change is evaluated by: (*a*) comparison of change score[10] variances for the 12 experimental groups and (*b*) comparison of differences between median change scores for treatments, diagnoses, and their interaction.

MMPI

The MMPI, like all experimental instruments, was administered when the patient arrived on the experimental ward and when he left the ward. In addition to the standard scales, the MMPI was scored for Barron's (1953) Ego Strength Scale (*Es*), Welsh's (1956) Anxiety (*A*) and Repression (*R*) scales, and Simon's (1952) Scale of Inner Maladjustment (*SIM*).

A comparison of the change score variances for the experimental groups on each scale was accomplished by use of Bartlett's test (Edwards, 1950).[11]

Heterogeneity of variance obtained on the *F, K, Mf, Pa, Pt, Sc, Es, SIM, R,* and *A* scales. Table 6 presents the mean variances for the 12 groups as well as the treatment and diagnostic groups.

A comparison of the mean variance for treatments, diagnoses, and their interaction is contained in Table 7.

TABLE 6

MEAN VARIANCES FOR THE 12 EXPERIMENTAL GROUPS ON CHANGE SCORES FOR THE 16 MMPI SCALES

	GG	G	I	C	Mean
N	148.56	74.66	128.56	91.08	110.72
Ps	152.89	281.87	216.68	65.89	179.30
Pl	351.76	188.78	189.72	194.88	231.28
Mean	217.74	181.77	178.32	117.28	

Table 7 shows that the mean variances for the 16 MMPI scales are significantly different between treatments, diagnoses, and their interaction. Table 6 reveals that all treatment groups have larger mean variance than the control with the group living treatment showing the greatest mean variance. For the diagnostic groups, short-term psychotics exhibit greater mean variance than non-psychotics and the long-term psychotic group shows the greatest variance. The higher variances in the short-term psychotic treatment groups compared with controls seems to account for the majority of the interaction variance.

Table 8 presents a comparison of median change scores by treatments, diagnoses, and their interaction. The treatment group comparisons and the interaction effects were evaluated by Friedman's test (Wilcoxon, 1949). The diagnostic group comparisons were made by the median test (Siegel, 1956). In these analyses, the only significant directional change between treatment groups occurred on the *Ma* scale. Here GG

TABLE 7

COMPARISON OF TREATMENTS AND DIAGNOSES ON 16 MMPI SCALE VARIANCES

Between Treatments χ_r^2 (3 *df*)	Between Diagnoses χ^2 (2 *df*)	Interaction: Treatment × Diagnoses χ_r^2 (6 *df*)
10.83*	15.15***	19.38**

* Significant at the .05 level.
** Significant at the .01 level.
*** Significant at the .001 level.

[9] There were no significant differences (.05 level) among the 12 experimental groups at initial testing on any of the assessment devices.

[10] Change scores refer to initial minus terminal test scores. Further use of the term "change score" refers to this difference.

[11] The change score variances for the 12 experimental groups, together with their tests of significance, on each scale presented in this study may be obtained from the authors.

and I groups showed gains in *Ma*, G group small loss, and C group a large loss. The only significant change between diagnoses occurred on Welsh's (1956) *R* scale. A reduction in *R* scores (repression) between initial and terminal testing occurred in the neurotic and short-term psychotic group, whereas an increase occurred in the long-term psychotic group.

Change scores on the *K*, *Pd*, and *SIM* scales yielded significant interactions. In all three significant interactions, the greatest share of the variance is attributable to the differential response of long-term psychotics in psychotherapy when compared with the controls. In these three interactions, one or all of the three long-term psychotic groups in psychotherapy demonstrated changes in the maladaptive scale direction while the controls remain relatively the same or change in the adaptive direction. On the other hand, nonpsychotics and short-term psychotics in psychotherapy change in an adaptive scale direction while the same diagnostic groups in the control condition changed only slightly or not at all.

Ward Behavior

Ward behavior is another area of adjustment often used as a treatment criterion. Because of its frequent use and because changes in ward behavior are often said to occur concomitantly with treatment, a Ward Behavior Scale pertinent to the experimental setting was constructed.

The scale was constructed in the following manner. Psychologists, social service personnel, nurses, and aides familiar with the experimental ward setting submitted behavior items which they believed were indicative of patients' adaptation to the ward.

Seventy-six of the 120 submitted items were selected by the experimenters for the Behavior Rating Scale. A review of the literature concerning behavior ratings revealed that 35 of the 76 selected items were so similar to Hospital Adjustment Scale (HAS) items (Ferguson, McReynolds, & Ballachey, 1953) that these 35 HAS items and their alternatives described below, were utilized as part of the Behavior Scale.

The 76 items were then grouped, a priori, into three areas of adaptation. They were (*a*) interpersonal communication—30 items, (*b*) self-care—31 items, and (*c*) participation in work and recreational activities—15 items.

The ratings were keyed in the following manner: for each of the 76 items, the experimenters

TABLE 8

COMPARISON OF TREATMENTS AND DIAGNOSES ON MMPI CHANGE SCORES

MMPI Scale	Between Treatments $\chi^2_r(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Diagnoses × Treatment $\chi^2_r(6\ df)$
1. *Lie*	3.23	3.53	10.87
2. *F*	1.21	0.08	4.25
3. *K*	0.64	0.75	18.75**
4. *Hs*	5.41	0.65	5.56
5. *D*	0.09	1.75	8.00
6. *Hy*	1.10	1.75	8.75
7. *Pd*	3.14	4.09	13.00*
8. *Mf*	2.43	0.59	10.06
9. *Pa*	2.34	1.62	4.81
10. *Pt*	2.34	3.59	8.50
11. *Sc*	1.51	0.75	4.75
12. *Ma*	10.21*	3.05	6.06
13. *Es*	0.14	1.00	10.25
14. *A*	0.94	1.09	8.00
15. *R*	1.06	12.19**	7.56
16. *SIM*	0.65	1.75	16.88**

* Significant at .05 level.
** Significant at .01 level.

made an a priori judgment as to whether the behavior described was adaptive or maladaptive in the ward setting. Then, four raters (one nurse and three aides) rated every patient one week after his admission to the experimental ward and again during his final week on the ward. Each item had two alternatives, one adaptive, the other maladaptive, i.e., "The patient usually will reply if you speak to him," or, "The patient usually does not reply if you speak to him." An item was scored in the following manner. If the four raters checked the adaptive alternative, the item was scored four; if three of the four raters checked the adaptive alternative, the item was scored three; if two raters checked the adaptive alternative, the score was two; if one rater checked the adaptive alternative, the score was one; if no rater checked the adaptive alternative, i.e., if all four checked the maladaptive, the score for the item was zero. A patient's total score was the sum of all items. His score for the three adjustment areas was the sum of the items for each area.

Comparison of the change score variances for the 12 experimental groups is presented in Table 9. Bartlett's (Edwards, 1950) test yields a χ^2 of 44.76 which, for 11 *df*, is significant beyond the .001 level. The marginal totals in Table 9 reveal that the greatest absolute behavioral change occurred in the group living therapy situation and, diagnostically, in the long-term psychotic groups.

A comparison of behavior change medians is presented in Table 10.

This table reveals that no significant differences in ward behavior median change scores obtain between treatments, diagnoses, or their interaction.

Holland Vocational Preference Inventory (HVPI)

It has been found by Holland (1958) and others that vocational preferences reflect personality characteristics. Changes in life adjustment could, theoretically, be revealed by such an inventory. The scale selected to evaluate this area was the Holland Vocational Preference Inventory (1958).

Holland's Inventory contains 300 items comprising 12 scales. These scales are briefly described as follows:

Scale 1. Physical Activity. This scale represents a concern with motor skills, physical strength, and a concrete approach to life problems.

Scale 2. Intellectuality. This scale measures preferences for conceptual methods of adjustment and strong needs for understanding the world and perhaps one's self. High scores may represent healthy ego functioning ("ego strength").

Scale 3. Responsibility. This scale appears to be an "oral receptive" scale. High scorers are dependent, verbal, and responsible.

Scale 4. Conformity. This scale is a measure of extraception in the sense of conforming, status-oriented, and conventional behavior.

Scale 5. Verbal Activity. This is an "oral aggressive" scale. High scorers are characterized as aggressive, oral, concerned with leadership and power.

Scale 6. Emotionality. This scale assesses the stability of the subject. High scorers are characterized as emotional and expressive. Low scorers are characterized as cool or "flattened."

Scale 7. Control. This scale measures the habitual inhibition of impulses to act out motives. High scorers are regarded as "over-controlled," and low scorers as "impulsive and under-controlled."

TABLE 9

VARIANCES FOR THE 12 EXPERIMENTAL GROUPS ON WARD BEHAVIOR CHANGE

	GG	G	I	C	Mean
N	1,157.00	757.43	319.86	1,108.14	835.61
Ps	888.57	188.29	828.00	399.71	576.14
Pl	7,713.42	4,101.00	1,601.42	1,105.00	3,630.21
Mean	3,253.00	1,682.24	916.43	870.95	

TABLE 10

COMPARISON OF TREATMENTS AND DIAGNOSES ON BEHAVIOR CHANGE

Ward Behavior Scale	Between Treatments $\chi^2_t(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Treatment × Diagnoses $\chi^2_t(6\ df)$
Total Behavior Rating	.86	1.00	4.25
Personal Care	4.19	0.08	4.25
Participation in Activities	1.00	3.41	1.25
Interpersonal Communication	.86	.58	5.56

Scale 8. Aggressiveness. This scale is a measure of aggressiveness with high scores representing aggressiveness and low scores passivity.

Scale 9. MF. This is an empirically constructed masculinity—femininity scale. High scores represent identification with or choice of typically masculine occupational roles; low scores represent the choice of feminine occupational roles.

Scale 10. Status. This scale represents the subject's need for status or prestige.

Scale 11. Infrequency. This scale is similar to the MMPI *F* scale. It is intended to provide an estimate of the degree to which the subject makes unusual or infrequent occupational choices.

Scale 12. Acquiescence. This scale indicates the response set of the subject, or the degree to which he responds "like."

Bartlett's test (Edwards, 1950) reveals that significantly different variances occur on the following scales: Conformity, Oral Aggression, Emotionality, *MF,* and Acquiescence.[12] The mean variances for the 12 experimental groups are presented in Table 11. Table 12 shows significant differences between mean variances for treatments and interaction. Table 11 shows the greatest mean variance in the three psychotherapy groups when compared to the control group. This table also reveals that the significant interaction is primarily due to an increase in variance of long-term psychotics in psychotherapy, whereas no such increase appears in the controls.

[12] Difference scores were computed in two ways. On those scales where the highest score is considered most adaptive, the difference scores were determined by substracting the initial from the terminal test score. On other scales, where the mean scale score is considered most adaptive, difference scores were determined by computing the unit movement toward or away from the mean.

TABLE 11

MEAN VARIANCES FOR THE 12 EXPERIMENTAL GROUPS ON CHANGE SCORES FOR THE 12 HVPI SCALES

	GG	G	I	C	Mean
N	24.43	24.87	48.67	38.51	34.12
Ps	35.95	33.12	47.50	13.94	32.63
Pl	53.49	82.22	104.78	19.02	64.88
Mean	37.96	46.74	66.98	23.82	

TABLE 12

COMPARISON OF TREATMENTS AND DIAGNOSES ON 12 HVPI SCALE VARIANCES

Between Treatments	Between Diagnoses	Interaction: Treatment × Diagnoses
$\chi^2_r(3\ df)$	$\chi^2(2\ df)$	$\chi^2_r(6\ df)$
30.25**	4.16	20.52*

*.01 level of significance.
**.001 level of significance.

Comparisons of treatments and diagnoses on median change are presented in Table 13.

Table 13 indicates that there are no significant differences between treatments or diagnoses. Two significant interactions occur—one on the intellectuality scale and the other on the masculinity–femininity scale. These two significant interactions, however, do not show any consistent differences when the median change scores for each of the 12 experimental groups are investigated. Due to the number of tests of significance

TABLE 13

COMPARISON OF TREATMENTS AND DIAGNOSES ON HVPI CHANGE SCORES

HVPI Scale	Between ments $\chi^2_r(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Treatment × noses $\chi^2_r(6\ df)$
1. Physical Activity	3.24	2.95	7.75
2. Intellectuality	1.71	.75	15.13*
3. Responsibility	2.71	2.18	9.32
4. Conformity	5.26	.25	5.26
5. Verbal Activity	3.35	1.62	10.25
6. Emotionality	1.69	.75	6.75
7. Control	5.08	.25	3.31
8. Aggressiveness	1.18	1.02	3.06
9. Masculinity-Femininity	5.09	.61	17.75**
10. Status	5.85	1.11	5.19
11. Infrequency	.13	.33	2.06
12. Acquiescence	2.65	1.09	7.63

* Significant at the .05 level.
** Significant at the .01 level.

computed, one should approach these findings with caution.

Q Sort

Q sort is another instrument frequently utilized to evaluate changes that occur during psychotherapy. In this study, a 100-item sort with five scales of 20 items each was used.

The general factor structure of Sheer's[13] Sort was followed; however, items appropriate to the experimental population were constructed. The five areas that the Q sort attempted to cover were: (*a*) social adaptability, (*b*) emotional control, (*c*) conformity, (*d*) self-acceptance, and (*e*) self-confidence. These were sorted both for self and ideal on initial and terminal testing. Each of the five scales contained 10 negative and 10 positive statements, the directionality of a statement being determined by the a priori judgment of the authors. For example, the self-confidence scale had as a positive item, "I can argue for my point of view" and, as a negative item, "I do as I am told." In scoring, positive items were given the score of the stack into which they were sorted,[14] and negative items received

[13] D. E. Sheer. A self-concept test: Reliability and validity data. Personal communication. 1955.

[14] For instructions and pile numbers, the reader is referred to the section on Experimental Instruments and Testing Situation.

TABLE 14

CORRELATIONS BETWEEN SELF AND IDEAL *Q* SORTS AT INITIAL AND TERMINAL TESTING

Subject	GG		G		I		C	
	Initial	Terminal	Initial	Terminal	Initial	Terminal	Initial	Terminal
	r	*r*	*r*	*r*	*r*	*r*	*r*	*r*
Nonpsychotic								
1	.67	.39	.45	.62	.53	.71	.32	.37
2	−.14	.39	.39	.42	.10	−.33	−.02	.03
3	−.35	−.12	−.51	−.43	.36	.60	.81	.79
4	.39	.61	.46	.42	.34	.49	.10	−.02
5	−.17	−.20	.78	.90	.14	.66	.64	.47
6	.52	.66	.28	.09	−.45	−.49	−.07	.09
7	.03	−.02	.54	.43	.18	−.27	.19	.58
8	−.48	−.31	−.25	.24	−.07	.15	.50	.34
Short-term Psychotic								
1	.34	.32	.21	.76	.50	.60	.09	−.19
2	.41	.33	−.14	.50	−.08	−.03	.06	.22
3	.66	.76	−.07	.09	.04	.05	.58	.64
4	−.11	.37	−.14	.02	.58	.73	.65	.65
5	.38	.68	.54	.61	.57	.69	.50	.46
6	−.09	.33	−.04	.21	−.18	.77	.73	.81
7	−.07	.00	.11	−.02	.56	.63	.51	.61
8	.46	.41	.11	.19	−.54	.65	.74	.81
Long-term Psychotic								
1	.32	−.08	.27	.19	−.16	.14	.78	.52
2	−.04	.37	−.08	.72	−.25	−.33	.58	.45
3	−.15	.43	−.05	.38	.25	−.23	.61	.58
4	.22	−.15	.53	.59	−.16	−.48	−.52	−.09
5	.23	.26	.12	.34	−.72	−.66	.07	.42
6	−.67	−.38	.65	.45	.58	.48	.87	.83
7	.31	−.08	.60	.29	.28	.54	.53	.61
8	.05	.10	.59	.24	.58	.03	.30	−.02

TABLE 15

VARIANCES FOR 12 EXPERIMENTAL GROUPS ON z SCORE CHANGE

	GG	G	I	C	Mean
N	.069	.074	.151	.064	.090
Ps	.069	.119	.286	.036	.128
Pl	.169	.229	.094	.117	.152
Mean	.102	.141	.177	.072	

the score of nine minus the pile number. The total score for each factor was the arithmetic sum of the 20 items in the scale. Thus, the following sets of measurements were obtained from the Q sort: (*a*) correlations between the 100-item self and ideal sorts at initial and terminal testing; (*b*) scores on self-sort on the five aforementioned scales at initial and terminal testing; (*c*) score on ideal-sort on the five scales at initial and terminal testing.

The usual comparisons made on Q sort data involve correlations between self- and ideal-sort at initial testing and between the same measures at terminal testing. Theoretically, successful therapy is evidenced by a positive increase in the correlations.

The correlations between self- and ideal-sorts at initial and terminal testing are presented in Table 14. These correlations were then transformed into z scores, and a population of differences between z scores (initial-terminal) was obtained.

The variances of this distribution for the 12 experimental groups is presented in Table 15. Barlett's test (Edwards, 1950) yields a χ^2 of 13.42 which, for 11 *df*, does not reach the .05 level of significance; however, the trend of larger variances in the three psychotherapy treatments and long-term psychotic diagnostic group occurs again.

TABLE 16

COMPARISON OF TREATMENTS AND DIAGNOSES ON z SCORE CHANGE

Between Treatments	Between Diagnoses	Interaction: Treatment × Diagnoses
χ^2_r(3 *df*)	χ^2(2 *df*)	χ^2_r(6 *df*)
2.85	4.00	4.50

TABLE 17

MEAN VARIANCES FOR THE 12 EXPERIMENTAL GROUPS ON CHANGE SCORES FOR THE 5 SELF-SORT SCALES

	GG	G	I	C	Mean
N	113.11	124.80	140.26	151.20	132.34
Ps	150.60	340.06	840.63	72.77	351.02
Pl	311.43	526.95	145.23	126.09	277.43
Mean	191.71	330.60	375.37	116.69	

Table 16 presents a comparison of treatments, diagnoses, and their interaction on the median differences between z scores. This table reveals that no significant differences obtain.

The previous comparisons dealt with self-ideal correlations based upon the 100 items. The following analyses are for the five Q sort scales of 20 items each. The first analysis compares change scores on self-sort alone. A Bartlett's test (Edwards, 1950) comparison of the self-sort change score variances showed significant differences between the variances on the social adaptability, self-acceptance, self-confidence, and emotional control scales.

The average variances for the five scales are presented in Table 17. A nonparametric analysis of the variances for the five scales is presented in Table 18 which shows that

TABLE 18

COMPARISON OF TREATMENTS AND DIAGNOSES ON 5 SELF-SORT SCALE VARIANCES

Between Treatments	Between Diagnoses	Interaction: Treatments × Diagnoses
χ^2_r(3 *df*)	χ^2(2 *df*)	χ^2_r(6 *df*)
4.28	5.20	22.80*

* Significant at the .001 level.

TABLE 19

COMPARISON OF TREATMENTS AND DIAGNOSES ON SELF-SORT CHANGE SCORES

Self-Sort Scale	Between Treatments $\chi^2_r(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Treatment × Diagnoses $\chi^2_r(6\ df)$
1. Emotional Control	4.85	.58	5.25
2. Conformity	.71	5.25	2.25
3. Self-Acceptance	3.18	2.58	15.75*
4. Social Adaptability	3.03	1.59	15.56*
5. Self-Confidence	3.29	.75	18.87**

* Significant at the .05 level.
** Significant at the .01 level.

the interaction alone is significant. Table 17 suggests that the interaction is significant because nonpsychotics show relatively equal variances over the four treatments, short-term psychotics show larger variances in the three groups receiving psychotherapy with the largest in the individual therapy group, and long-term psychotics show the largest variances in the group living and group therapy treatments.

Table 19 presents a comparison of the medians for the self-sort change scores. No significant differences obtain on any of the five scales between treatments or diagnoses. However, three significant interactions obtain.

When the three significant interactions are viewed collectively, it is evident that the major contributor to the significant interaction is the differential response of the diagnostic groups to all three psychotherapy treatments compared with the control. The following schema generally holds for the three scales:

	All Therapies	Control
Nonpsychotics	Large Positive Change	No Change or Small Positive Change
Short-term Psychotics	Large Positive Change	Small Positive or Small Negative Change
Long-term Psychotics	Moderate Negative Change	Large Positive Change

TABLE 20

MEAN VARIANCES FOR THE 12 EXPERIMENTAL GROUPS ON CHANGE SCORES FOR THE 5 IDEAL-SORT SCALES

	GG	G	I	C	Mean
N	86.94	47.37	138.42	140.77	103.38
Ps	95.80	85.67	237.24	49.16	116.97
Pl	200.91	174.72	162.95	95.55	158.53
Mean	127.88	102.59	179.54	95.16	

As was previously mentioned, initial and terminal ideal-sorts were also obtained. These were scored on the five scales.

A comparison of the ideal-sort change score variances for the 12 experimental groups revealed significantly different variances on four of the five scales: emotional control, conformity, self-acceptance, and self-confidence.

The average of the five scale variances for the 12 experimental groups is shown in Table 20. An analysis of the five variances when treated as scores appears in Table 21. Only significant differences between the variances of the treatment groups obtain. Table 20 indicates that the variances in descending order of magnitude are: individual therapy, group living, group therapy, and control.

An analysis of the median change scores on ideal-sort is presented in Table 22. Significantly different changes occurred between diagnostic groups on the emotional control, conformity, and social adaptability scales. The short-term psychotic group has the largest median improvement in emo-

TABLE 21

COMPARISON OF TREATMENTS AND DIAGNOSES ON THE 5 IDEAL-SORT SCALE VARIANCES

Between Treatments $\chi^2_r(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Treatments × Diagnoses $\chi^2_r(6\ df)$
9.66*	2.80	10.40

* Significant at the .05 level.

TABLE 22

COMPARISON OF TREATMENTS AND DIAGNOSES ON IDEAL-SORT CHANGE SCORES

Ideal-Sort Scale	Between Treatments $\chi^2_r(3\,df)$	Between Diagnoses $\chi^2(2\,df)$	Interaction: Treatment × Diagnoses $\chi^2_r(6\,df)$
1. Emotional Control	2.60	6.59*	3.94
2. Conformity	7.00	7.60*	5.50
3. Self-Acceptance	2.88	1.58	5.25
4. Social Adaptability	2.16	6.78*	5.06
5. Self-Confidence	1.24	1.59	8.75

*.05 level of significance.

tional control and conformity, whereas both the non- and short-term psychotic groups demonstrate greater median improvement in the area of social adaptability than do the long-term psychotics.

Thematic Apperception Test

Eight selected TAT cards were presented to the subjects at the initial and terminal testing sessions. The cards were selected by the authors for their value in evoking responses concerning interpersonal relationships. In order of presentation they were: 1, 2, 4, 9, 13 MF, 7 BM, 6 BM, and 12 M.

Protocols were scored in the following manner: All identifying information was first removed. Then, the initial and final protocols for each patient were presented to the experimenter who made a judgment as to whether one protocol was more adaptive or whether they were adaptively equivalent. The scoring system for the protocols was:

I. Protocol A is judged more adaptive than Protocol B when Protocol B has four or more rejected cards and Protocol A has none.

II. Protocol A is judged more adaptive than Protocol B when A's stories reveal developed plots and B's are descriptions of the cards.

III. Protocol A is judged more adaptive than B if both reveal plots but A's are more well-developed than B's.

IV. Protocol A is judged more adaptive than B when both reveal well-developed plots but A's stories contain direct affective statements, i.e., love, hate, etc., while B's do not.

V. Protocol A is judged more adaptive than B when both reveal well-developed plots, both contain direct affective statements, but Protocol A has more positive statements concerning interpersonal relationships than B.

VI. Protocol A is judged more adaptive than B when both reveal well-developed plots, both contain direct affective statements, both have an equivalent number of positive statements concerning interpersonal relationships, but A's outcome is more positive than B's outcome.

After the judgments were completed, identifying information was revealed and scores were recorded. The scores were recorded by comparing final with initial protocols as follows: final more adaptive than initial, adaptively equivalent, or less adaptive than initial.

Two experimenters independently compared the initial and final protocols of the same 20 patients by this scoring system. They agreed in 18 of the 20 (or 90%) of the cases.

For statistical analysis the scores were dichotomized into protocols showing posi-

TABLE 23

TAT CHANGES FOR THE 12 EXPERIMENTAL CONDITIONS BY %

		GG	G	I	C	Total %
N	Positive change	75.0	50.0	12.5	62.5	50.0
	No change and negative change	25.0	50.0	87.5	37.5	50.0
Ps	Positive change	37.5	75.0	37.5	0.0	37.5
	No change and negative change	62.5	25.0	62.5	100.0	62.5
Pl	Positive change	25.0	50.0	75.0	50.0	50.0
	No change and negative change	75.0	50.0	25.0	50.0	50.0
	Total %					
	Positive change	45.8	58.3	41.7	37.5	
	No change and negative change	54.2	41.7	58.3	62.5	

tive change and those showing no change or negative change. Table 23 presents the percentage of cases in the two categories for each of the 12 cells.

Sheffield's (1957) method was used to compare the treatment and diagnostic groups. Table 24 presents the analysis.[15]

Table 24 indicates that differences between treatments and diagnoses do not reach the .05 level of significance. The interaction is significant at the .01 level. Reference to Table 23, however, suggests that this interaction is significant because more positive TAT change is brought about by group methods and control than with individual therapy when nonpsychotics are treated; more positive TAT change occurs with the three psychotherapy groups than controls when short-term psychotics are treated; and more positive TAT change is brought about by individual therapy when long-term psychotics are treated.

TABLE 24

COMPARISON OF TREATMENTS AND DIAGNOSES ON TAT CHANGE

	Squares	*df*	*MS*	*F*
Treatments	4.67	3	1.56	0.53
Diagnoses	2.67	2	1.34	0.67
Interaction: Treatment × Diagnoses	35.33	6	6.06	3.06**
Total	42.67	11	3.88	1.96*

*.05 level of significance.
**.01 level of significance.

Follow-up

A follow-up questionnaire was obtained for each patient six months after he had left the experimental program. The questionnaire was completed by an individual with whom the patient was living, or, if he had not left or had returned to the hospital, by a psychologist familiar with the case.[16]

Follow-up Scale. The questionnaire was constructed to encompass several areas of life adjustment. Simplicity of items was stressed. The items were multiple choice, with a designated space for remarks if the respondent believed the structured choices to be inadequate. The questionnaire consisted of five objective and four subjective items. The five objective items concerned: (*a*) percentage of time employed, (*b*) hospitalization, (*c*) alcoholic indulgence, (*d*) antisocial behavior, and (*e*) number of friends. The four subjective items involved a judgment by the respondent concerning patient's (*a*) amount of verbal communication with others, (*b*) general adjustment, (*c*) degree of problem behavior, and (*d*) degree of illness.

For analysis, the nine items were divided into the following dichotomies:

1. Was, or was not, rehospitalized
2. Was, or was not, employed full-time
3. Has some, or no, close friends
4. Frequently, or rarely, talks to the people with whom he lives
5. Getting along satisfactorily, or poorly
6. Usually hard, or easy, to get along with
7. Well, or sick, most of the time
8. Is, or is not, a problem drinker
9. Has, or has not, been in trouble with the law

Each of the nine items, therefore, has an arbitrary adaptive and maladaptive alternative. The tetrachoric correlations among the nine follow-up items are presented in Table 25.

Table 25 reveals that, in general, positive relationships exist among the items. The high degree of relationship among Items 5, 6, and 7 indicates a common parameter and the theoretical possibility of item substitution. However, the remaining relationships, while generally positive and often significant, are rarely of sufficient magnitude to warrant the substitution of one item for another.

Comparisons of the follow-up items by treatments, diagnoses, and their interactions, computed by Sheffield's (1957) method, are presented in Table 26.

Table 26 indicates that the amount of employment item is the only one that significantly differentiates the treatment groups. The percentage of patients from each treatment group employed full-time is as fol-

[15] The within-cell variance is 1.98.

[16] Information was obtained on all 96 patients for Items 1 and 2. On Items 3 through 9, information was obtained on 94 of the 96 subjects. The scores of the two missing patients was the modal score for his experimental group; however, deletion of these two patients would not alter the results. Eleven of the 96 patients were failures, i.e., did not leave the hospital. Viewed diagnostically, nine were in the long-term psychotic group and two in the nonpsychotic group. Viewed by treatments, four were in group living (GG), two in group therapy (G), and five in individual therapy (I).

TABLE 25

CORRELATIONS AMONG THE 9 FOLLOW-UP ITEMS

	1	2	3	4	5	6	7	8	9
1. Rehospitalization									
2. Full-time employment	.66**								
3. Number of friendships	.12	.46*							
4. Amount of interpersonal communication	.35	.45	.78**						
5. General adjustment	.56*	.40	.54*	.55*					
6. Degree of problem behavior	.42	.42	.78**	.36	.90**				
7. Degree of illness	.59**	.57*	.50*	.36	.93**	.91**			
8. Type of alcoholic indulgence	.29	.14	−.08	.11	.77**	.59*	.60*		
9. Legal violations	.65**	.25	−.22	−.22	.71*	.62*	.44	.75**	

* Significant at .01 level.
** Significant at .001 level.

lows: Group living, 42%; Group therapy, 33%; Individual therapy, 42%; and Control, 8%. It can readily be seen that the three treatments involving psychotherapy have a higher percentage of employed than the control and that the group living and individual therapy treatments have the highest percentage of the three treatments. Table 1 indicates that this result cannot be explained by differences in employment prior to the study.

Diagnostically, six of the nine items significantly differentiate the groups. Table 27 presents the percentage of the three diagnostic groups in the adaptive alternative for each item. This table indicates that in the areas of hospitalization, employment, and friendships, the short-term psychotic group displayed the most adequate adjustment, the nonpsychotics next, and the long-term psychotics the poorest adjustment. In the two areas involving judgments concerning

TABLE 26

COMPARISON OF EXPERIMENTAL GROUPS ON FOLLOW-UP ITEMS

Follow-Up Area	Between Treatments *F* ratio (3 *df*)	Between Diagnoses *F* ratio (2 *df*)	Interaction: Treatment × Diagnoses *F* ratio (6 *df*)
1. Rehospitalization	0.16	8.16**	1.02
2. Amount of employment	2.78*	6.25**	0.50
3. Number of friendships	1.26	4.81**	1.60
4. Amount of interpersonal communication	0.56	2.18	1.07
5. General adjustment	0.55	3.05*	0.44
6. Degree of problem behavior	2.08	0.89	0.78
7. Degree of illness	0.66	3.89*	0.81
8. Type of alcoholic indulgence	0.83	4.56*	1.33
9. Legal violations	0.15	2.12	0.92

Note.—The within-cell variance is 1.98.
* Significant at the .05 level.
** Significant at the .01 level

TABLE 27

% OF DIAGNOSTIC GROUPS SHOWING BEST ADAPTATION ON THE 6 SIGNIFICANT FOLLOW-UP ITEMS

Follow-up Area	Non-Psychotic	Short-Term Psychotic	Long-Term Psychotic
1. Out of hospital	50.0	87.5	40.6
2. Employed full-time	34.4	50.0	9.4
3. Some friends	62.5	75.0	37.5
5. Getting along satisfactorily	53.1	81.3	59.4
7. Well most of the time	53.1	84.4	59.4
8. Not a problem drinker	65.6	90.6	90.6

general adjustment (Items 5 and 7), again the short-term psychotics were rated as making a significantly better adjustment than either the nonpsychotics or long-term psychotics. Problem drinkers are more prevalent in the nonpsychotic group.

Relationships among Selected Psychotherapeutic Criteria

Thirteen scales chosen because they are frequently used as psychotherapeutic criteria and because they measure adjustment in different areas were selected for correlational analysis. These scales were: follow-up measures of Rehospitalization, Employment, and Degree of Illness; *Sc, Es, SIM, A,* and *R* scales of the MMPI; Q sort; TAT; Ward Behavior; Job Behavior;[17] and Finney's (1954) Group Therapy Scale.

Only the last two scales will be described here since the others were discussed in preceding sections. The Job Behavior Scale included 31 items dealing with the patient's performance and interpersonal adjustment in his hospital job assignment. These items, like those in the Ward Behavior Scale, were stated in terms of adaptive and maladaptive alternatives; i.e., "The patient usually profits by his mistakes" or "The patient repeats his mistakes over and over again." The rating was completed by each patient's job supervisor in the first and final week of the experimental program. A reliability coefficient of .93 was found for 20 patients, by the method of successive ratings. Each patient in group therapy was rated on the Group Therapy Scale every 30 days.

All scales were scored so that the most adaptive scores were numerically highest and the least adaptive were lowest. Thus, only positive correlations indicate an association between the theoretical adaptive levels of the scales. For all within-hospital measures, final scores completed shortly before discharge were correlated. In the case of the Group Therapy Rating Scale, the last rating was used. For the Q sort, the final correlations between self- and ideal-sort were ranked, with the highest positive correlation receiving the highest rank. The final 96 TATs were divided into two groups —most and least adaptive.[18]

The *N* for all scales was 96 except for Job Behavior with an *N* of 70 and the Therapy Rating Scale which had an *N* of 41. All scales, except the follow-up items, were divided at the median for computational purposes.

The correlations are presented in Table 28. An examination of the matrix reveals two rather independent clusters of items. One involves the three follow-up items and the group therapy behavior rating. The other cluster includes the *Sc, A, Es,* and *SIM* scales of the MMPI, the Q sort, and the TAT as a marginal member.

Three of the scales are not contained in these two clusters. They are Job Performance, Ward Behavior, and Welsh's (1956) *R*. Hospital job behavior is not significantly related to any other measure and Welsh's *R* is correlated only with ward behavior. On the other hand, ward behavior is significantly related only to the *R* scale, group therapy behavior, and TAT.

That test scores are not significantly related to follow-up and are associated with

[17] The Job Behavior Scale has been deposited with the American Documentation Institute. See Footnote 8.

[18] General principles of scoring outlined in the TAT section were followed.

each other is not surprising (Fairweather, Moran, & Morton, 1956). But why is group therapy behavior the only within-hospital measure significantly related to follow-up? The group therapy rating completed by the therapist after the patient's last session was used for the correlational analysis. Thus, this rating involves continuing observations of an individual over several weeks' time. Further, in contrast to ward behavior and job performance ratings, the group therapy ratings were made by psychologists rather than by nurses, aides, and job supervisors. In addition, the group therapy situation involves intense interpersonal experiences which may be a better approximation of the extramural situation than either ward behavior or the hospital job situation. Therefore, both the rater and the situation may have contributed to these significant relationships.

Discussion

The purpose of this study was to test a series of hypotheses in an attempt to determine the relative efficiency of four commonly utilized therapeutic procedures with three different patient groups. In view of the lack of agreement concerning the validity of a specific criterion for evaluating treatment, several criteria were selected.

In neuropsychiatric hospitals, economic considerations are important in determining the feasibility of therapeutic endeavors because most hospitals operate on limited and frequently minimal budgets. Therefore, the amount of time taken to complete a specified program is a relevant factor in the selection of treatment programs. The analysis of time in treatment yielded significant differences, both for treatment and diagnostic groups. Individual therapy patients remained in treatment longer than the other three groups. Group living took less time, group therapy and control the least, and approximately equivalent, time. By diagnostic group, the long-term psychotic remained in treatment significantly longer than did the short-term and nonpsychotic groups. In terms of time alone, it appears that group therapy and individual work assignment are the most economical. Diagnostically, nonpsychotics or short-term psychotics can be treated most economically (shortest time in hospital).

While time in treatment is an important consideration in the selection of psycho-

TABLE 28

Tetrachoric Correlations among Therapeutic Criteria

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Follow-up Rehospitalization													
2. Follow-up Employment	.66[c]												
3. Follow-up Degree of Illness	.59[b]	.57[b]											
4. Finney's Group Therapy Behavior	.42[a]	.75[c]	.09										
5. Ward Behavior	.04	.28	−.32	.50[a]									
6. Job Behavior	.02	−.06	.08	−.14	−.14								
7. MMPI Welsh's *R*	.06	.05	.07	.29	.41[b]	−.24							
8. MMPI *Es*	.30	.15	−.01	.12	.07	−.19	−.10						
9. MMPI Welsh's *A*	.30	.18	.16	.05	.07	.30	−.16	.75[c]					
10. MMPI *Sc*	.21	.02	.29	.11	.23	−.08	.20	.53[b]	.58[c]				
11. MMPI *SIM*	.26	.09	−.09	.12	−.04	−.27	.26	.78[c]	.74[c]	.58[c]			
12. *Q* Sort	.21	.10	−.14	.20	.10	−.29	.07	.64[c]	.75[c]	.66[c]	.87[c]		
13. TAT	.04	.23	−.21	.41[a]	.33[a]	−.27	.10	.13	.20	.26	.16	.47[b]	

[a].05 level of significance.
[b].01 level of significance.
[c].001 level of significance.

therapeutic programs, it sheds little light upon the success or failure of such programs. In an attempt to ascertain the relationship between time in treatment and successful follow-up, a "global follow-up" score[19] was computed for each patient. A comparison of treatments, diagnoses, and their interaction on this criterion is presented in Table 29. The only significant difference is between diagnoses; hence, comparisons of success and time in treatment were independently computed for the three diagnostic groups. Table 30 shows no significant differences in mean time in treatment for the least and most adaptive follow-up groups.[20] Thus, length of time in treatment appears to have little relationship to successful posthospital adjustment. This result is reasonable if we consider psychotherapy as a learning process. Such a process is unique to each individual and each learns at his own rate. If such is the case, the amount learned may be unrelated to time in the learning situation. Perhaps future studies might profitably attempt to discriminate slow and fast learners prior to treatment.

[19] All nine follow-up items were dichotomized into adaptive and maladaptive alternatives. The adaptive alternative received a score of one and the maladaptive a score of zero. Each patient's score was the sum of the nine items.

[20] Each diagnostic group was divided at its median global follow-up score. Least and most adaptive refer to patients below and above this median.

TABLE 29

COMPARISON OF TREATMENTS AND DIAGNOSES ON GLOBAL FOLLOW-UP SCORES

Between Treatments $\chi^2_r(3\ df)$	Between Diagnoses $\chi^2(2\ df)$	Interaction: Treatment × Diagnoses $\chi^2_r(6\ df)$
2.15	21.75*	4.38

* Significant at .001 level.

Change was evaluated in the preceding sections by analyzing the difference scores between initial and final testing on the MMPI, behavior ratings, HVPI, TAT, and Q sort. The first finding was that significant differences in variances obtain between treatment and diagnostic groups. When change scores are viewed globally, the generic conclusion, that the three treatments involving psychotherapy demonstrate significantly more absolute change than the control, seems warranted. These differences do not appear to be simply a function of time between tests. Table 9 shows that the individual treatment group presents approximately the same variance as the controls despite the fact that these two groups spent the greatest and least time in treatment, respectively. Further, absolute change seems to be related to type of therapy involved. For example, the greatest variance in ward behavior is achieved by the group

TABLE 30

COMPARISON OF MEAN DAYS IN TREATMENT FOR PATIENTS IN LEAST AND MOST ADAPTIVE FOLLOW-UP CATEGORIES

	Diagnoses					
	N		Ps		Pl	
	Most Adaptive	Least Adaptive	Most Adaptive	Least Adaptive	Most Adaptive	Least Adaptive
Mean	117.57	81.25	89.35	90.58	160.56	147.71
SD	65.35	45.35	71.18	75.35	74.22	57.55
t	1.79		.04		.54	

living subjects, whereas the greatest Q sort variance is achieved by the individual therapy group. Group living therapy was oriented toward interpersonal behavior in a variety of situations, while individual therapy was more closely oriented toward perceptual change. Thus, the different variances of the treatment groups in different areas of change might be, in some measure, an indication of the attendant responses in the learning situation.

While significantly larger variances were found for treatment groups as compared to controls, significantly different variances also occurred between diagnostic groups. Generally, significantly larger variances obtained for the long-term psychotics when compared with the short-term and nonpsychotic groups. Again, it seems highly unlikely that the markedly different variances are attributable to time alone, although clearly they are confounded with time. From observation of the patients in the treatment processes, it seems likely that the significantly greater change score variances of the long-term psychotic group can be attributed, in part, to the greater range of responses available to them. In this regard, the repertoire of responses available to the long-term psychotic in reacting to a close interpersonal relationship ranges from, to take one example, complete withdrawal to hyperactive participation. By comparison, the range of responses available to the nonpsychotic is much more restricted and, on this theoretical continuum, the short-term psychotics would fall somewhere between these two groups. In addition, long term psychotics in the three psychotherapy groups contribute greatly to the total long-term psychotic group variance, while the controls contribute little. Thus, as far as the scales are concerned, the long-term psychotics, and to a lesser but usually significant extent, the short-term psychotics, change in either a positive or negative scale direction in psychotherapy while the same diagnostic groups remain comparatively unchanged without psychotherapy.

There are, of course, several possible explanations for the different variances. One is that patients in therapy get better or worse but generally do not remain the same, whereas patients not in therapy remain about the same. A second is inherent in the construction of the scales themselves. Most clinical instruments are constructed on the assumption of an underlying theoretical dimension of adjustment. Thus, positive change on a scale usually refers to an increment of movement along the scale in the adaptive direction. In this context, subjects show improvement only by moving in the same direction. Clinically, however, this assumption may not be tenable. For example, two patients bearing the same diagnostic label, one a rigid conformist and the other a nonconformist, might be expected, if therapy was successful, to move in opposite directions. One might move 20 positive scale units and the other 20 negative units on a Conformity Scale. The mean of the two changes is zero. Two untreated patients might move two positive and two negative units. Here again, the mean is zero, and the difference between means is obviously not significant. The variances, however, of the two groups would reveal the change. That movement in either direction on a scale may be an adaptive change for the individual patient is emphasized in

TABLE 31

COMPARISON OF SELECTED CHANGE SCORES FOR PATIENTS IN LEAST AND MOST ADAPTIVE FOLLOW-UP CATEGORIES

Scale	Diagnoses		
	N	Ps	Pl
	χ^2	χ^2	χ^2
Job Behavior	.05	.39	.88
Ward Behavior	.13	.00	.00
MMPI *F*	.98	.68	.00
MMPI *K*	.00	.01	2.17
MMPI *Sc*	.00	.00	.45
MMPI *Es*	6.22*	1.20	1.14
MMPI *A*	.00	.01	.04
MMPI *SIM*	.13	.68	.13
HVPI 2	.98	.01	.45
HVPI 3	.02	.01	.00
Q Sort	1.14	.15	.00

* Significant at .05 level.

two additional analyses. Table 31 shows that only one significantly greater positive change occurs for the least and most adaptive follow-up groups, and this change was in the opposite direction predicted, i.e., least adaptive on follow-up showed greater positive scale change. On the other hand, Table 32 reveals significantly different variances for the least and most adaptive follow-up groups on five scales for the short-term psychotics and three for the long-term psychotics. Furthermore, on each significantly different scale, the larger variance was associated with the most adaptive follow-up. Thus, it seems that change in either scale direction can be adaptive for the individual patient. For this reason, it might be clinically more meaningful if a directional prediction of adaptive change were made for each patient prior to therapy. Thus, positive change for a patient would be determined by the direction in which clinicians agree that the patient should move rather than by the scale itself.

Although generally the variances of the change scores significantly differentiated the treatment groups, differences between change score medians did not. Of the within-hospital measures, only the *Ma* scale of the MMPI showed a significant difference and, in view of the large number of measures, it seems plausible that little significance should be attached to this finding.

Interactions, on the other hand, revealed significance on nine scales. On six of the nine significant interactions, the three psychotherapy groups showed change in an adaptive direction for the short-term and nonpsychotic groups, while an adaptive decrement occurred with the long-term psychotics. Generally, nonpsychotics and short-term psychotics in the control group remained relatively unchanged, while long-term psychotics showed a small or large adaptive increment. As far as scale score adaptiveness is concerned, then, nonpsychotics and short-term psychotics change more adaptively with psychotherapy while long-term psychotics change more adaptively without therapy.

On within-hospital measures, the diagnostic groups showed significantly different

TABLE 32

COMPARISON OF SELECTED CHANGE SCORE VARIANCES FOR PATIENTS IN LEAST AND MOST ADAPTIVE FOLLOW-UP CATEGORIES

	Diagnoses								
	N			Ps			Pl		
Scale	Most Adaptive s^2	Least Adaptive s^2	F	Most Adaptive s^2	Least Adaptive s^2	F	Most Adaptive s^2	Least Adaptive s^2	F
Job Behavior	9.80	31.81	3.25	20.25	11.00	1.84	113.61	8.98	12.65*
Ward Behavior	388.25	934.02	2.41	735.30	186.22	3.95*	2218.31	2122.67	1.05
MMPI *F*	85.63	45.14	1.90	197.41	30.91	6.39*	340.17	331.23	1.03
MMPI *K*	32.12	52.69	1.64	74.29	52.06	1.43	254.10	52.39	4.85*
MMPI *Sc*	309.39	145.89	2.12	397.41	369.74	1.28	305.58	635.84	2.09
MMPI *Es*	51.35	36.40	1.41	89.54	32.85	2.73	182.31	47.03	3.88*
MMPI *A*	66.39	53.92	1.23	174.29	16.72	10.42*	167.53	64.23	2.61
HVPI 2	11.67	34.56	2.96	59.15	26.92	2.20	80.11	101.35	1.27
HVPI 3	29.39	37.57	1.28	110.03	26.39	4.17*	115.06	86.29	1.33
MMPI *SIM*	262.37	324.25	1.24	510.95	195.75	2.61	1142.02	549.54	2.08
Q sort	.09	.06	1.50	.16	.04	4.00*	.09	.20	2.22

* Significant at .05 level.

change on four scales. Although such findings must be interpreted cautiously, they suggest, as set forth in the second hypothesis, that the short-term and nonpsychotic groups improve more with treatment than long-term psychotics.

Follow-up items showed significant differences between treatment groups on the employment item. All three treatment groups showed a greater percentage of full-time employed than the controls, which supports the first hypothesis. The group living and individual therapy groups have the highest percentages of employed patients. The eight other items did not significantly differentiate the treatment groups. If employment were chosen as the criterion of posthospital adjustment, then either the group living or group therapy treatments are desirable because they take less time and personnel than individual therapy and are significantly more successful than the control. Further, such therapy would be more efficient with nonpsychotics and short-term psychotics than with long-term psychotics. On the other hand, if any other follow-up item besides employment were chosen as the criterion, e.g., rehospitalization, then hospital work alone would be the most efficient treatment. In this regard, it should be noted that follow-up employment is the only item yielding any measure of the patient's contribution to others. It will be recalled that nine significant interactions occur on the within-hospital measurements, but no significant interaction is found on follow-up. Thus, differential test score change of the diagnostic groups to treatment is not revealed by the follow-up.

Diagnostically, six of the nine items significantly differentiated the groups; the short-term psychotics demonstrate the most adequate global posthospital adjustment, nonpsychotics are next, and long-term psychotics have the poorest adjustment. These findings again support the second hypothesis. The significant differences on follow-up ratings indicate all too clearly the problem facing treatment planners. The long-term psychotic, in this study, is much less responsive to psychotherapy than the nonpsychotic or short-term psychotic.

In addition, Table 33 indicates that patients who did not receive tranquilizing medication did as well on follow-up as patients who did. Further, the table reveals that the results were the same for all three diagnostic groups. Since tranquilizing medication is frequently used to improve ward behavior, i.e., reduce management problems, and since ward behavior is unrelated to follow-up, it is not surprising that the patients who did not receive drugs did as well on follow-up as those who did.

It is interesting to note that nine of the eleven failures in the total program were long-term psychotics in psychotherapy. No failures occurred in the control group. On the other hand, the three full-time employed long-term psychotics were in psychotherapy and none in the control. This finding suggests that long-term psychotics in intensive

TABLE 33

COMPARISON OF PATIENTS WHO RECEIVED TRANQUILIZING MEDICATION IN LEAST AND MOST ADAPTIVE FOLLOW-UP CATEGORIES

Diagnoses									
N			Ps			Pl		
	Most Adaptive	Least Adaptive		Most Adaptive	Least Adaptive		Most Adaptive	Least Adaptive
Drugs	6	6	Drugs	11	7	Drugs	11	10
No Drugs	14	18	No Drugs	20	12	No Drugs	18	14
	$\chi^2 = .30$			$\chi^2 = .03$			$\chi^2 = .15$	

interpersonal situations change for better or worse while avoiding such situations results in stabilization at a marginal adjustment level.

The correlations among follow-up items are, in general, positive and significant. Yet it seems quite clear that the correlations are not of sufficient magnitude to warrant the assumption of a single criterion. Thus, the outcome of any treatment program is a function of the follow-up criterion chosen, which is based upon a value judgment. For instance, one would arrive at different conclusions concerning the efficacy of treatment programs in the current study if one arbitrarily decided that the criterion of hospitalization was more important than the criterion of employment.

But further complications in the criteria problem are revealed in the correlation matrix shown in Table 28. Within-hospital measures, with the exception of the group therapy rating, are not significantly related to follow-up. Thus, ward behavior, long an accepted criterion in treatment evaluation, is not significantly related to the three follow-up items chosen and performance in the hospital work situation is not significantly related to any other measure. Tests form a cluster of interrelated scales which are, by and large, not significantly related to other criteria. The relationships between follow-up criteria and the Group Therapy Rating Scale suggests that observations by trained personnel of intensive interpersonal situations within the hospital may be the best predictors of follow-up adjustment.

The fact that there are two relatively independent criteria clusters, (*a*) subjective self-evaluation as revealed in tests and (*b*) objective interpersonal evaluation as revealed in follow-up items and group therapy rating, caution against generalizing from test measures to within-hospital behavior or follow-up behavior. Implications of these relationships for treatment evaluation studies using a single criterion are obvious.

SUMMARY

This investigation attempted to evaluate the relative efficiency of four psychotherapeutic programs with three patient groups on a number of different criteria. All patients in the study lived on the same experimental ward. The four treatments were: (*a*) group living, in which all activities were group oriented, i.e., work, living arrangements, psychotherapy, etc., (*b*) group therapy and an individual work assignment, (*c*) individual therapy and an individual work assignment, and (*d*) a control group with an individual work assignment and no therapy. The three patient groups were: (*a*) a nonpsychotic group, (*b*) a short-term psychotic group, and (*c*) a long-term psychotic group. The design was a three by four factorial type analysis of variance with an N of eight in each cell and a total N of 96. All patients were tested and rated upon entry into the program and upon departure. The instruments were: (*a*) Ward Behavior, (*b*) MMPI, (*c*) TAT, (*d*) Holland Vocational Preference Inventory, (*e*) Q sort, and (*f*) a six-month follow-up questionnaire. In addition, all patients in group therapy were rated every 30 days on Finney's (1954) Therapy Scale and 70 of the 96 were rated on job performance, both upon entry and termination of the program.

The following results obtained:

1. The number of days in treatment was significantly different for both treatment and diagnostic groups. The individual therapy group took the greatest time, group living next, with group therapy and control the least and almost equivalent times. Diagnostically, long-term psychotics were in treatment longest, with short-term psychotics and nonpsychotics requiring the least and almost equivalent times.

2. Change score comparisons (initial–terminal test scores) were computed for variances and medians. Generally, significantly different variances occurred on most instruments between treatments and diagnoses. The control group usually had the smallest variance and the three psychotherapy groups the largest. Of the four treatments, group living or individual therapy usually showed the greatest variance. Variances were usually significantly

different between diagnoses with the long-term psychotics having the largest variances, followed in descending order by the short-term and nonpsychotic patients.

3. Tests of significance between change score medians revealed only the *Ma* scale of the MMPI yielding a significant difference between treatments. However, nine interactions were significant, and six of these followed the same pattern: nonpsychotics and short-term psychotics in the three groups receiving psychotherapy showed moderate to large adaptive change while long-term psychotics in the same three groups showed moderate to large change in a maladaptive scale direction. On the other hand, nonpsychotics and short-term psychotics in the control group showed relatively little change, and long-term psychotics in the same program showed moderate to large adaptive scale score change. Four scales showed significantly different median change between the diagnostic groups. On all of these scales the nonpsychotic and short-term psychotic groups changed more adaptively than the long-term psychotics.

4. Correlations between the nine follow-up items revealed three that were so highly intercorrelated as to be considered a single criterion. Other items were, in general, significantly correlated but not of sufficient magnitude to warrant criteria substitution. Tests of significance between treatments revealed significant differences in employment, with the three groups which had received psychotherapy having a higher percent of full-time employed than the control. The other eight follow-up items revealed no significant differences between treatments. When compared by diagnoses, six of the nine follow-up items significantly differentiated the groups. Overall, the short-term psychotics made the most adequate posthospital adjustments, nonpsychotics next, and the long-term psychotics the poorest.

5. Thirteen frequently used scales were selected for correlational analysis. They were: three follow-up items, (Rehospitalization, Employment, Degree of Illness), three within-hospital behavior measures, (Ward Behavior, Job Performance, and Group Therapy Behavior), TAT, Q sort, and *Sc*, *Es*, *A*, *R*, and *SIM* from the MMPI. Two relatively independent clusters of interrelated measures appeared: (*a*) three follow-up items and group therapy behavior and (*b*) MMPI scales of *Sc*, *Es*, *A*, *SIM*, Q sort, and TAT as a marginal member.

6. Other comparisons revealed that "global" follow-up adjustment was unrelated to time in treatment or receipt of tranquilizing drugs during treatment.

The results were discussed in relation to treatment programs, psychotherapy criteria, and further research.

REFERENCES

BARRON, F. An ego-strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, **17**, 327–333.

EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950. Pp. 195–197.

FAIRWEATHER, G. W., MORAN, L. J., & MORTON, R. B. Efficiency of attitudes, fantasies, and life history data in predicting observed behavior. *J. consult. Psychol.*, 1956, **20**, 58.

FERGUSON, J., MCREYNOLDS, P., & BALLACHEY, E. *The hospital adjustment scale.* Stanford, Calif.: Stanford Univer. Press, 1953.

FINNEY, B. C. A scale to measure inter-personal relationships in group psychotherapy. *Group Psychother.*, 1954, **7**, 52–66.

HOLLAND, J. L. *Manual for the HVPI.* Palo Alto, Calif.: Consulting Psychologists Press, 1958.

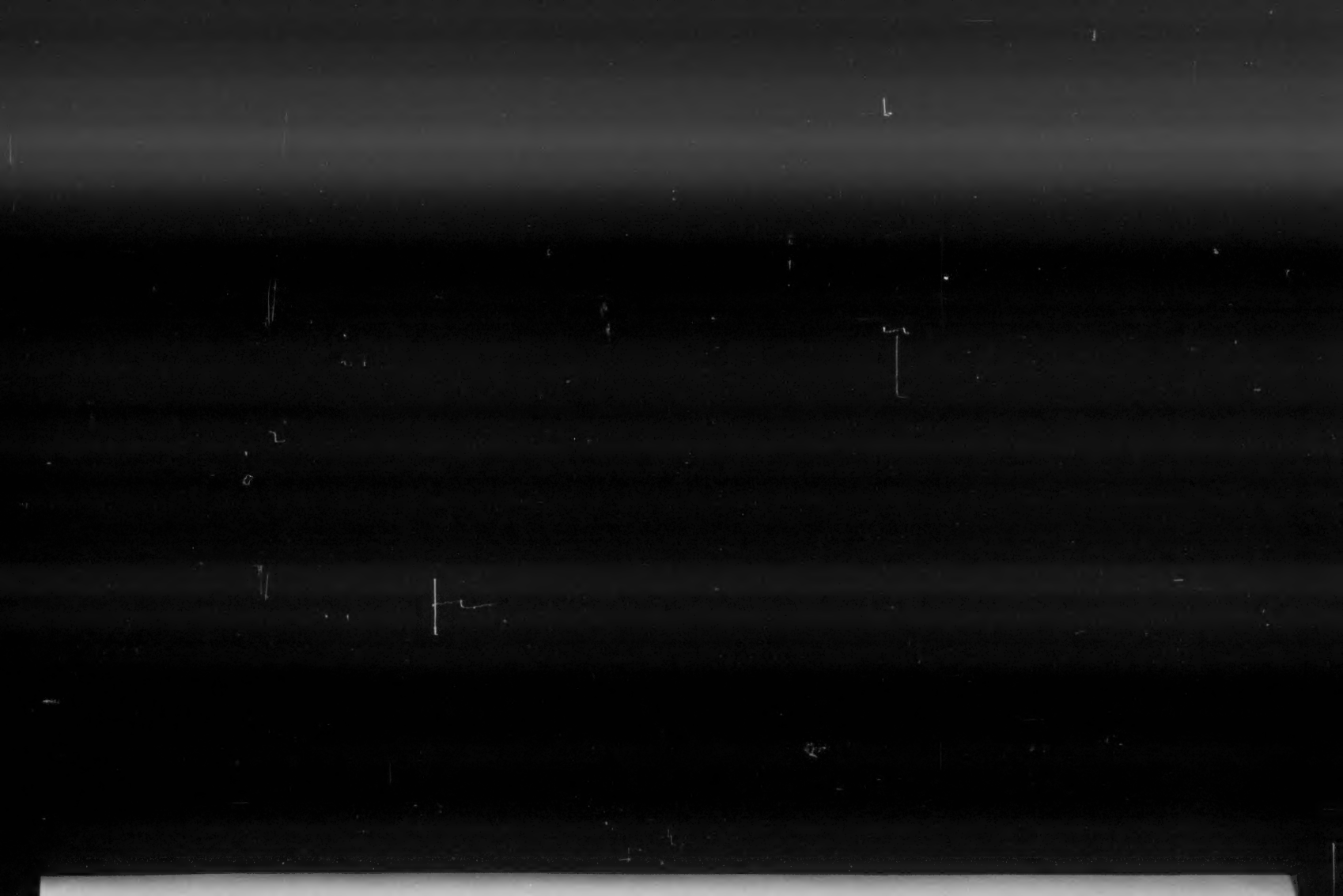
SHEER, D. E. A self-concept test: Reliability and validity data. Personal communication, 1955.

SHEFFIELD, F. D. Comment on a distribution-free factorial-design analysis. *Psychol. Bull.*, 1957, **54**, 426–428.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956. Pp. 179–184.

SIMON, R. An investigation of the relationship between several measures of self-insight and adjustment, intelligence and personality characteristics. Unpublished doctoral dissertation, Syracuse Univer., 1952.

WELSH, G. S. Factor dimensions *A* and *R*. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine.* Minneapolis, Minn.: Univer. Minnesota Press, 1956. Pp. 264–281.

WILCOXON, F. *Some rapid approximate statistical procedures.* New York: American Cyanamid, 1949. P. 5.

(Received March 24, 1959)